

LogisticRegression.R

Administrator

Tue Apr 05 15:02:14 2016

```
# logistic regression
# use when the response variable is discrete (0,1)
# predictor variable is continuous
```

```
# create ordered x values
xdat<- sort(runif(20,min=0,max=100))
```

```
# create y response data yes/no
ydat<- c(0,0,0,0,0,1,1,0,1,1,1,0,1,1,1,0,1,1,1)
```

```
# bind in a data frame
MyData <-data.frame(cbind(xdat,ydat))
```

```
# inspect data
head(MyData)
```

```
##          xdat ydat
## 1  0.3204023    0
## 2  8.1686429    0
## 3 13.0676654    0
## 4 19.7528953    0
## 5 20.9191853    0
## 6 22.1992966    1
```

```
# fit logistic regression model
MyModel<-glm(ydat~xdat, family = binomial("logit"), data=MyData)
summary(MyModel)
```

```
##
## Call:
## glm(formula = ydat ~ xdat, family = binomial("logit"), data = MyData)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0960  -0.7537   0.4043   0.7300   1.4817
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.95900    1.11546  -1.756   0.0791 .
## xdat         0.05708    0.02534   2.252   0.0243 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
## Null deviance: 26.920 on 19 degrees of freedom
## Residual deviance: 19.452 on 18 degrees of freedom
## AIC: 23.452
##
## Number of Fisher Scoring iterations: 4
```

```
# create empty plot with grid
plot (x=xdat,y=ydat,xlab="Plant Height", ylab="P(pollination)",type="n")
grid()

# add curve of predicted values
curve(predict(MyModel,data.frame(xdat=x),type="resp"),add=TRUE)

# add predicted points
points(x=xdat,y=fitted(MyModel),pch=20)

# add observed points
points(x=xdat,y=ydat,cex=2,pch=21,bg="lightblue")
```

